

**Amendment and Response**

Applicant: Parvathi Chundi et al.  
Serial No.: 10/767,151  
Filed: January 29, 2004  
Docket No.: 10990670-2  
Title: DOCUMENT CLUSTERING METHOD AND SYSTEM

**REMARKS**

The following remarks are made in response to the Office Action mailed May 16, 2006. Claims 17-35 were rejected. With this Response, claims 26, 29, 30, and 32 have been amended. Claims 17-35 remain pending in the application and are presented for reconsideration and allowance.

**Claim Objections**

The Examiner objected to claims 29 and 30 because both of these claims refer to claim 9, which was cancelled.

Claims 29 and 30 have been amended to depend from independent claim 28. Accordingly, Applicants submit that the above objection to claims 29 and 30 should be withdrawn.

In addition, claim 26 has been amended to correct a typographical error and claim 32 has been amended to depend from independent claim 31.

**Claim Rejections under 35 U.S.C. § 103**

The Examiner rejected claims 17-22 under 35 U.S.C. § 103(a) as being unpatentable over Cooley et al, "WebSIFT: The Web Site Information Filter System," (1999) ("Cooley").

Applicants submit that Cooley fails to teach or suggest the invention recited by independent claim 17 including performing log-based clustering on the session logs to generate session clusters; representing each session cluster as a log-based document suitable for content based clustering; receiving a plurality of documents that includes a first document that was accessed in one session and a second document that was not accessed in the sessions; replacing the first document with a log-based document associated with the session cluster that includes the first document; and performing content based clustering on at least the first document and the second document to generate clusters with user perspective.

Cooley discloses applying data mining techniques to large Web data repositories to extract usage patterns. (Abstract). The Web Site Information Filter system is a Web Usage Mining framework that, in addition to performing preprocessing and knowledge discovery, uses the structure and content information about a Web site to automatically define a belief

**Amendment and Response**

Applicant: Parvathi Chundi et al.

Serial No.: 10/767,15

Filed: January 29, 2004

Docket No.: 10990670-2

Title: DOCUMENT CLUSTERING METHOD AND SYSTEM

set. The information filter uses this belief set to identify results that are potentially interesting. (§ 1, Introduction and Background).

In addition to failing to address each of the particular limitations of claim 17 listed above, the Examiner admits that Cooley does not explicitly teach about log-based or content based clustering. (Office Action, page 3). Despite the substantial differences between claim 17 and the disclosure of Cooley, the Examiner nonetheless rejected claim 17 under 35 U.S.C. § 103(a). Since the Examiner did not cite any other references in rejecting claim 17, the Examiner appears to be relying on Official Notice. As indicated in the Manual of Patent Examining Procedure, however, “[o]fficial notice unsupported by documentary evidence should only be taken by the examiner where the facts asserted to be well-known, or to be common knowledge in the art are capable of instant and unquestionable demonstration as being well-known.” M.P.E.P. § 2144.03(A). “It would not be appropriate for the examiner to take official notice of facts without citing a prior art reference where the facts asserted to be well known are not capable of instant and unquestionable demonstration as being well known.” *Id.* (emphasis in original). Applicants contend that the limitations of claim 17 that the Examiner appeared to indicate were not disclosed by Cooley are not well known facts that are capable of instant and unquestionable demonstration as being well-known. Applicants respectfully request allowance of this claim, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach each limitation of claim 17.

In view of the above, Applicants respectfully submit that the above rejection of independent claim 17 under 35 U.S.C. § 103(a) should be withdrawn. Dependent claims 18-22 further define patentably distinct independent claim 17. Accordingly, Applicants believe that these dependent claims are also allowable over the cited reference. Allowance of claims 18-22 is respectfully requested.

In addition, Cooley fails to teach or suggest the invention recited by dependent claim 18 including wherein representing each session cluster as a log-based document suitable for content based clustering includes modifying each document referenced in the session cluster so that a Euclidean distance between the documents is the same. The Examiner admits that Cooley does not explicitly teach how to make the Euclidean Distance between documents the same. (Office Action, page 3). Again, since the Examiner did not cite any references regarding a Euclidean distance, the Examiner appears to be relying on Official

**Amendment and Response**

Applicant: Parvathi Chundi et al.

Serial No.: 10/767,151

Filed: January 29, 2004

Docket No.: 10990670-2

**Title: DOCUMENT CLUSTERING METHOD AND SYSTEM**

Notice. Accordingly, Applicants respectfully request allowance of this claim, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach the further limitations of claim 18.

The Examiner rejected claims 23-27 under 35 U.S.C. § 103(a) as being unpatentable over Cutting et al., "Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections," (1993) ("Cutting").

Applicants submit that Cutting fails to teach or suggest the invention recited by independent claim 23 including generating a hybrid matrix of vectors comprising a first vector representing a first document and a second vector representing a log-based document cluster; and clustering the documents using the hybrid matrix.

Cutting discloses a document browsing method called Scatter/Gather, which uses document clustering as its primitive operation. The technique is directed towards information access with non-specific goals and serves as a complement to more focused techniques. (§1, Introduction). In the basic iteration of the browsing method the user is presented with short summaries of a small number of document groups. Initially the system scatters the collection into a small number of document groups, or clusters, and presents short summaries of them to the user. Based on these summaries, the user selects one or more of the groups for further study. The selected groups are gathered together to form a subcollection. The system then applies clustering again to scatter the new subcollection into a small number of document groups, which are again presented to the user. With each successive iteration the groups become smaller, and therefore more detailed. Ultimately, when the groups become small enough, the process bottoms out by enumerating individual documents. (§2, Scatter/Gather Browsing).

In addition to failing to address each of the particular limitations of claim 23 listed above, the Examiner admits that Cutting does not teach the specific method of constructing the hybrid matrix. (Office Action, page 4). Again, since the Examiner did not cite a reference teaching each of the limitations of claim 23, the Examiner appears to be relying on Official Notice. Accordingly, Applicants respectfully request allowance of this claim, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach each limitation of claim 23.

**Amendment and Response**

Applicant: Parvathi Chundi et al.

Serial No.: 10/767,151

Filed: January 29, 2004

Docket No.: 10990670-2

Title: DOCUMENT CLUSTERING METHOD AND SYSTEM

In view of the above, Applicants respectfully submit that the above rejection of independent claim 23 under 35 U.S.C. § 103(a) should be withdrawn. Dependent claims 24-27 further define patentably distinct independent claim 23. Accordingly, Applicants believe that these dependent claims are also allowable over the cited reference.

In addition, Cutting fails to teach or suggest wherein a second vector is used in place of a second document within the hybrid matrix wherein the second document forms a portion of the log-based document cluster as recited in dependent claim 24, wherein clustering the documents using the hybrid matrix is performed using a content-based clustering technique as recited in dependent claim 25, wherein generating the hybrid matrix comprises: accessing retrieval session logs; clustering retrieval sessions into session clusters; generating a log-based document cluster for each session cluster by combining all documents opened during any retrieval session of the session cluster; generating a log-based document cluster vector for each of the log-based document clusters; replacing each document in the log-based document cluster with the log-based document cluster vector; generating an individual document vector for each document not opened during any retrieval session; and combining the log-based document cluster vector and the individual document cluster vector as recited in dependent claim 26, and wherein the step of clustering retrieval sessions into session clusters comprises the steps of: generating a Boolean session vector for each retrieval session; forming a matrix of the Boolean session vectors; and applying a clustering algorithm to the matrix of the Boolean session vectors as recited in dependent claim 27.

The Examiner admits that Cutting does not teach the specific clustering algorithm as read in claims 24-27. (Office Action, page 4). Again, since the Examiner did not cite any references teaching the further limitations of claims 24-27, the Examiner appears to be relying on Official Notice. Accordingly, Applicants respectfully request allowance of these claims, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach the further limitations of claims 24-27.

The Examiner rejected claims 28 and 35 under 35 U.S.C. § 103(a) as being unpatentable over Cooley in view of Pitkow et al., U.S. Patent No. 6,457,028 ("Pitkow").

Pitkow merely discloses a computer system including a processor and memory. (Col. 12, lines 12-14; Fig. 10).

**Amendment and Response**

Applicant: Parvathi Chundi et al.

Serial No.: 10/767,151

Filed: January 29, 2004

Docket No.: 10990670-2

Title: DOCUMENT CLUSTERING METHOD AND SYSTEM

For the same reasons as discussed above with reference to claim 17, Cooley and Pitkow, either alone, or in combination, fail to teach or suggest the invention recited by independent claim 28 including a processor connected to the storage, configured to cluster the retrieval sessions into session clusters, generate, for each session cluster, a log-based document cluster, generate a log-based document cluster vector for each of the log-based document clusters, generate an individual document vector for each document not opened during any retrieval session, cluster the documents using the log-based document cluster vectors and individual document vectors.

In view of the above, Applicants respectfully submit that the above rejection of independent claim 28 under 35 U.S.C. § 103(a) should be withdrawn. Accordingly, like claim 17, Applicants respectfully request allowance of this claim, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach each limitation of claim 28.

For the same reasons as discussed above with reference to claim 17, Cooley and Pitkow, either alone, or in combination, fail to teach or suggest the invention recited by independent claim 35 including the data structure having entries for a log-based document cluster vector generated from a log-based document cluster, and an individual document vector corresponding to a vector generated from a first document, the first document not belonging to any log based document cluster.

In view of the above, Applicants respectfully submit that the above rejection of independent claim 35 under 35 U.S.C. § 103(a) should be withdrawn. Accordingly, like claim 17, Applicants respectfully request allowance of this claim, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach each limitation of claim 35.

The Examiner rejected claims 29 and 30 under 35 U.S.C. § 103(a) as being unpatentable over Cooley in view of Pitkow, and in further view of Cutting. Dependent claims 29 and 30 further define patentably distinct independent claim 28. Accordingly, Applicants believe that these dependent claims are also allowable over the cited references. Allowance of claims 29 and 30 is respectfully requested.

The Examiner rejected claims 31-34 under 35 U.S.C. § 103(a) as being unpatentable over Pitkow in view of Cooley.

For the same reasons as discussed above with reference to claim 17, Pitkow and Cooley, either alone, or in combination, fail to teach or suggest the invention recited by

**Amendment and Response**

Applicant: Parvathi Chundi et al.

Serial No.: 10/767,151

Filed: January 29, 2004

Docket No.: 10990670-2

Title: DOCUMENT CLUSTERING METHOD AND SYSTEM

independent claim 31 including a media readable by the processor having a document clustering module having a plurality of instructions, that when executed by the processor, performs log-based clustering on the session logs to generate session clusters, converts the session clusters into a form suitable for content-based clusters, performs content-based clustering on the documents and session clusters in a form suitable for content-based clustering to generate document clusters with users' perspective.

In view of the above, Applicants respectfully submit that the above rejection of independent claim 31 under 35 U.S.C. § 103(a) should be withdrawn. Accordingly, like claim 17, Applicants respectfully request allowance of this claim, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach each limitation of claim 31.

Dependent claims 32-34 further define patentably distinct independent claim 31. Accordingly, Applicants believes that these dependent claims are also allowable over the cited references. Allowance of claims 32-34 is respectfully requested.

In addition, Pitkow and Cooley, either alone, or in combination, fail to teach or suggest wherein the document clustering module further comprises: a session vector generation module for receiving the session logs and based thereon for generating a session vector for each session log; a session cluster generation module coupled to the session vector generation module for receiving the session vectors and based thereon for generating session clusters; a hybrid matrix builder for receiving the documents, coupled to the session cluster generation module, for receiving the session clusters and based thereon for generating a hybrid matrix having at least one log-based document; and a topic generation module coupled to the hybrid matrix builder for receiving the hybrid matrix and based thereon for generating document clusters with users' perspective as recited in dependent claim 32, and wherein the hybrid matrix builder further comprises: a session document generation module for receiving session clusters and based thereon generates super documents; and document modification module coupled to the session document generation module for receiving the super documents, for receiving the documents, and based thereon for generating the hybrid matrix as recited in dependent claim 33.

Again, since the Examiner did not cite any references teaching the further limitations of claims 32 and 33, the Examiner appears to be relying on Official Notice. Accordingly,

**Amendment and Response**

Applicant: Parvarhi Chundi et al.

Serial No.: 10/767,151

Filed: January 29, 2004

Docket No.: 10990670-2

**Title: DOCUMENT CLUSTERING METHOD AND SYSTEM**

---

Applicants respectfully request allowance of these claims, or request pursuant to M.P.E.P. § 2144.03 that the Examiner cite a reference to teach the further limitations of claims 32 and 33.

**Amendment and Response**

Applicant: Parvathi Chundi et al.  
Serial No.: 10/767,151  
Filed: January 29, 2004  
Docket No.: 10990670-2

Title: DOCUMENT CLUSTERING METHOD AND SYSTEM

---

**CONCLUSION**

In view of the above, Applicants respectfully submit that pending claims 17-35 are in form for allowance and are not taught or suggested by the cited references. Therefore, reconsideration and withdrawal of the rejections and allowance of claims 17-35 is respectfully requested.

No fees are required under 37 C.F.R. 1.16(h)(i). However, if such fees are required, the Patent Office is hereby authorized to charge Deposit Account No. 08-2025.

The Examiner is invited to contact the Applicant's representative at the below-listed telephone numbers to facilitate prosecution of this application.

Any inquiry regarding this Amendment and Response should be directed to either Steven E. Dicke at Telephone No. (612) 573-2002, Facsimile No. (612) 573-2005 or Lloyd E. Dakin, Jr. at Telephone No. (650) 857-2295, Facsimile No. (650) 852-8063. In addition, all correspondence should continue to be directed to the following address:

IP Administration  
Legal Department, M/S 35  
HEWLETT-PACKARD COMPANY  
P.O. Box 272400  
Fort Collins, Colorado 80527-2400

Respectfully submitted,

Parvathi Chundi, et al,

By their attorneys,

DICKE, BILLIG & CZAJA, PLLC  
Fifth Street Towers, Suite 2250  
100 South Fifth Street  
Minneapolis, MN 55402  
Telephone: (612) 573-2002  
Facsimile: (612) 573-2005

Date: September 18, 2006  
SED:uw

Steven E. Dicke  
Steven E. Dicke  
Reg. No. 38,431